Subjective Cues to Deception/Honesty in a High Stakes Situation: An Exploratory Approach

Clea Wright Whelan[1], Graham F. Wagstaff[2] and Jacqueline M. Wheatcroft[2]

Address correspondence to Clea Wright Whelan, Department of Psychology, University of Chester, Parkgate Road, Chester, CH1 4BJ, UK; c.wrightwhelan@chester.ac.uk (email)

[1] Department of Psychology, University of Chester, Chester, United Kingdom

[2] Department of Psychological Sciences, University of Liverpool, Liverpool, United Kingdom

This work was supported by the Economic and Social Research Council [ES/I90316X/1]

Abstract

The low ecological validity of much of the research on deception detection is a limitation recognised by researchers in the field. Consequently, the present studies investigated subjective cues to deception using the real life, high stakes situation of people making public appeals for help with missing or murdered relatives. It was expected that cues related to affect would be particularly salient in this context. Study 1 was a qualitative investigation identifying cues to deception reportedly used by people accurate at detecting deception. Studies 2 and 3 were then empirical investigations which mainly employed the cues reported in Study 1. A number of subjective cues were found to discriminate between honest and deceptive appeals, including some previously unidentified cues, and cues likely to be context-specific. Most could be categorised under the themes of authenticity of emotion, and negative and positive affective reactions to the appealer. It is concluded that some cues to deception may emerge only in real life, high stakes situations; however, it is argued that some of these may be influenced by observers' perceptions of the characteristics of offenders, rather than acts of deception per se.

.

Key words: Deception; Lie Detection; High Stakes Lies

Subjective Cues to Deception in a High Stakes Situation: Public Appeals for Help with Missing or Murdered Relatives

The low ecological validity of the majority of deception research is a limitation that has been recognised by a number of researchers in the field (see, for example, Frank & Svetieva, 2012; Granhag & Stromwall, 2004; Porter & ten Brinke, 2010). So far, most deception research has been confined to low stakes, laboratory studies, the findings of which may be applicable to the low stakes situations in which they were carried out, but may have less relevance to more high stakes situations, such as most forensic contexts. In particular, the use of low stakes situations may lead to underestimates of the effectiveness of some cues in detecting deception in more realistic forensic situations. For example, in their meta-analysis of cues to deception, DePaulo et al. (2003) found that strong motivation to succeed in the lie, and lies about transgressions, factors more likely to be present in high stakes situations, emerged as important moderating factors in the elicitation of prominent cues to deception. Indeed, the small body of existing research investigating real life, high stakes lies, has produced some promising results, and suggests that there may be some useful cues to deception that are more likely to emerge in high stakes situations; for example, an increase in speech dysfluency (Davis, Markus, Walters, Vorus & Connors, 2005; Vrij & Mann, 2001a), and equivocal language (Adams & Jarvis, 2006; Wright Whelan, Wagstaff & Wheatcroft 2013).

However, DePaulo and Morris (2004) have suggested that there may be further, previously unidentified, cues to deception that may emerge only in high stakes situations, and may be specific to particular forensic contexts (see also, Porter & ten Brinke, 2010). For example, in their study of 911 homicide calls, Harpster, Adams and Jarvis (2009) found 19 behaviours that discriminated significantly between innocent callers and callers who were later found to be involved in the homicides they were reporting. Importantly, a number of these behaviours had not been previously identified and were likely to be specific to the

particular context; for example, acceptance of death, plea for the caller only, and insulting the victim. Furthermore, it can be noted that a number of cues identified in this study are not readily accommodated by the four major factors that are popularly considered by researchers to underlie the production of cues to deception, as summarised in Zuckerman, DePaulo and Rosenthal's (1981) Four Factor Model: Lying involves complex and simultaneous cognitive tasks, and consequently liars are expected to produce behaviours associated with increased cognitive load. Liars are also expected to experience affective responses associated with the act of deception, typically, shame, guilt, fear and anxiety. Liars may take their credibility for granted less than truth-tellers, and so may control their behaviour in an attempt to suppress signs of deception, and also to simulate credible behaviours. And the act of deception may produce psychological discomfort, resulting in increased arousal in the autonomic nervous system, and associated involuntary physiological reactions. The inability of these existing theoretical approaches to adequately account for the newly identified cues in Harpster et al.'s study (2009), suggests that new theoretical approaches may be required to account for cues that may emerge only in specific, high stakes contexts.

To maximise the possibility of finding such previously unidentified cues, it is also important to consider the methodology employed. For instance, despite meta-analytic findings that suggest that subjective impressions and implicit measures may be more powerful than objective measures in discriminating between honesty and deception (DePaulo et al., 2003), the focus in most of the research so far on high stakes lies has mainly been on objective cues that allow investigation with frequency counts; i.e. few have examined more subjective cues to deception (see, for example, Koper & Sahlman, 1991; Mann & Vrij, 2006). Given these considerations, the overall aim of the present research was to investigate the possible presence of cues to deception and to honesty using subjective observer ratings of

behaviours, in a specific high stakes situation; that is, public appeals for help with regard to missing or murdered relatives.

Three studies are described, all of which employed real-life video clips of people appearing in front of the press, appealing for help with regard to a missing or murdered relative. Some of these were deceptive (i.e. the appealer was later found to be involved in the death or disappearance of their relative), and some were truthful (the appealer was not involved in the death or disappearance of their relative). The first study was a qualitative investigation of the subjective cues to deception and to honesty reportedly used by a small group of people found to be unusually accurate at detecting deception in appeals. The purpose of this study was to identify possible cues that may not have previously been identified and may be specific to the context of making appeals. Thus an exploratory, inductive approach was used, rather than theoretical prediction. The aim of the second study was to investigate which of the cues identified in the first study successfully discriminated between honest and deceptive appealers using the same set of stimulus materials, but using a different, independent panel of raters. Those cues found to be useful in discriminating between honest and deceptive appeals in Study 2 were then included in a third validation study, which used an entirely new and larger stimulus set. As the initial study was essentially inductive and exploratory, no explicit hypotheses were proposed at the outset; however, because appeals for missing or murdered relatives are likely to be an emotionally charged event, it was assumed that affective responses (i.e. responses related to grief, anxiety etc.) might be particularly salient.

Study 1

To reiterate, the aim of this first study was to generate subjective cues to deception and to honesty, reportedly used by a small group of people found to be unusually accurate at detecting deception in appeals.

Method

Participants. Two participants were recruited for the study; both had taken part in a previous study on cues to deception, and had achieved accuracy rates above 87% in distinguishing between deceptive and honest appealers (see Wright Whelan, 2009).

Materials and procedure. The study used nine video clips of people making appeals for help with missing or murdered relatives; five involved honest appealers (another person has been convicted of the murder or kidnap of the relative), and four, deceptive appealers (the appealer has been convicted of murdering their relative). The two participants were asked, individually, to view each appeal with regard to whether it was deceptive or honest, and to report and comment on any behaviours they used to decide if the appealer was deceptive or truthful. After viewing each clip a semi-structured interview schedule was used to prompt the participants to comment further on aspects which they may not previously have mentioned.

Data Analysis and Discussion

The responses of the two participants were analysed using the inductive thematic analysis procedure described by Hayes (2000) and Braun and Clarke (2006). The data were analysed as two separate elements, deceptive appealers and honest appealers. First, the responses were read carefully and repeatedly to identify units of text relevant to the research topic, and to search for meanings and patterns. Second, units relating to the same topic were grouped together into categories and initial codes were generated. Units of text could be included in more than one category, and the entire content of the data set was coded. The data were

reviewed to ensure that each category was sufficiently supported by the units. The analysis identified five key themes for the deceptive appealers; fake emotion, lack of emotion, distancing, implausibility, and negative personal reaction. Four key themes were identified for the honest appealers; genuine sadness, genuine/heartfelt, containing emotion, and sympathetic personal reaction. Each key theme had a number of associated categories, resulting in a total of 50 cues. The themes and categories, with the number of units in each category, are presented in Table 1.

It is relevant to note here that the conceptual distinction between cues to honesty and cues to deception is an important one. The premise underlying this distinction is that while the presence of a particular cue may probabilistically indicate honesty, its absence does not necessarily indicate dishonesty; similarly while the presence of a particular cue may probabilistically indicate deceit, its absence does not necessarily indicate honesty. Indeed, to infer one from the other would be to commit a logical fallacy of the form, 'if A then B, therefore, if not A then not B'.

It can be noted that although some of these cues have been investigated elsewhere, the majority have not previously been identified. This suggests that the type of inductive approach used in this study may be useful for identifying potential cues that have not been suggested by existing theoretical approaches. Amongst these, cues related to the affect of the appealers, and the affective responses of the observers to the appealers, are perhaps of particular interest, and their possible theoretical significance is considered later. However, the first task was to establish whether cues of this kind could be used to reliably discriminate between truthful and deceptive appeals.

Study 2

As noted previously, Study 2 was an empirical investigation to validate the cues to deception and to honesty identified in Study 1, using the same materials but with a different,

independent panel of raters. Since the sample size was small, and consequently there was a risk of Type II errors in analyses of statistical significance, it was decided to also examine effect sizes in order to illuminate meaningful differences between deceptive and honest appeals (Howell, 2002).

Method

Participants. Five new observers were recruited on a voluntary basis to rate the materials. Ages of the raters ranged from 21 to 75. One rater was a postgraduate psychology student and four were non-psychologists. None had previously received any formal training in lie detection techniques.

Materials and procedure. The video clips were the same as those used in Study 1. Observers were shown an appeal, and were then required to complete a response sheet, which contained one statement relating to each cue identified in Study 1 (for example, 'This person's emotions are fake'). Observers were asked to indicate their agreement with each statement on a five point Likert scale from 1 (strongly disagree) to 5 (strongly agree). This process was repeated for each clip. The clips were presented in a different, randomised order for each observer, and each observer completed the rating procedure individually.

Results and Discussion

To assess inter-rater reliability between the five observers, Kendall's W was calculated for each cue. Inter-rater reliability was not significant (p>.05) for seven of the cues, which were dropped from further analysis (calm and calculated, lack of emotion in voice, irrelevant/generic statements, positive emotion expressed towards relative, smile appropriate to speech content, avoid brutal language/detail, and behaving as expected). Average Spearman r for dropped cues ranged from r = .011 (smile appropriate to speech content) to r

= .230 (irrelevant/generic statements). Average Spearman *r* for cues that were retained ranged from *r* = .256 (creepy) to *r* = .663 (pretending to cry). The mean score for each cue in each clip was calculated from the scores of the five observers, so that each clip had a single score for each cue; these scores were used in all further analyses.

To investigate which cues were associated with deception, a series of one way ANOVAs was conducted on the means for each cue between the deceptive appealers and the honest appealers (veracity condition). Notwithstanding the small sample sizes (5 and 4), at this stage the main purpose of the study was still exploratory (to identify cues worthy of further investigation) hence the parametric ANOVA was chosen as it is a relatively powerful test, yet fairly robust to violations of the standard assumptions with fixed levels of the independent variable and fairly equal sample sizes (Glass & Stanley, 1970). A series of ANOVAs was also chosen rather than a MANOVA, as there were more dependent variables than cases per cell, thus violating the minimal sample size requirement for MANOVA (Tabachnick & Fidell, 2001).

As shown in Table 2, cues which yielded significant effects between liars and truth-tellers were as follows: fake emotion, fake facial expression, putting on a performance, creepy, no sympathy for the appealer, high vocal pitch, gaze aversion, and listing things/people, were related to deceptive appeals. Sad eyes, containing emotion, and genuine facial expression were related to honest appeals. Two cues only very narrowly missed significance on a two-tailed test ($p<.06$), and had means in the expected directions; these were, credible eye contact and genuinely sad. Again, in view of the exploratory nature of the study, no correction for multiple testing was applied.

As previously mentioned, it was considered useful to examine effect sizes as well; according to Cohen (1988), effect sizes of 0.8 can be considered as large, and a number of

cues showed effect sizes of this magnitude ($d \geq 0.80$) in the predicted directions. These were, behaving unnaturally, dislike of the appealer, pretending to cry, no emotional variation, emotionally cold, equivocal language, focusing on others, defensive, and does not make sense, all of which were associated with deceptive appeals. Genuine/heartfelt appeal, feeling sorry for the appealer, normal, feeling the appealer's pain, and genuine behaviours, were associated with honest appeals.

Levene's test for homogeneity of variance was significant for the cues fake facial expression, no sympathy for appealer, equivocal language, focus on others, genuine facial expression and feel sorry for the appealer. Welch's Test was run on these cues as a check, and the results were comparable to the parametric tests (fake facial expression $p = .034$, no sympathy for appealer $p = .041$, equivocal language $p = .093$, focus on others $p = .192$, genuine facial expression $p = .047$, and feel sorry for the appealer $p = .113$).

In sum, 13 cues were found to discriminate significantly, or very near significantly, between honest and deceptive appealers, with a further 14 cues having large effect sizes in the expected direction. These findings suggest that a large number of behaviours identified by the accurate participants in Study 1 may potentially have some utility in discriminating between deceptive and honest appealers in. Nevertheless, given the small sample sizes used, and the very marginal nature of many of the statistical effects found, there is obviously merit in testing further the reliability of these trends on a new and larger sample set. This was the purpose of Study 3.

Study 3

Given the issues concerning the robustness of the findings in Study 2, the aim of Study 3 was to replicate and extend the trends identified in Study 2 on a new and larger sample set. To this end, three further subjective cues identified in previous research were included; urgency (Harpster et al., 2009), general plausibility (Koper & Sahlman, 1991; DePaulo et al., 2003),

and a personal and expressive voice (DePaulo et al., 2003). Also included was one cue noted by the researcher (voice quivering with genuine emotion). A small number of verbal cues were dropped as other research has indicated that they may be more effectively investigated using frequency counts; these were, lists, equivocal language, focusing on others, and statement that does not make sense (Wright Whelan et al., 2013). Another cue, genuine facial expression, was also dropped as it was considered to be the negative equivalent, and, therefore, repetition, of fake facial expression. It can be noted that actuarial rather than theoretical prediction provided the primary rationale for the inclusion of cues at this stage; possible theoretical explanations for their efficacy are explored later. Due to the lack of previous research in this area, and also for reasons discussed previously, it was decided to include non-significant cues with large effect sizes in the expected direction, to minimise the risk of Type II errors occurring.

Method

Participants. Five observers were used to rate the appeals, all were undergraduate psychology students participating for course credit. Again, none had previously received any formal training in lie detection techniques.

Materials and procedure. The procedure was the same as for Study 2, except observers were shown 32 video clips of appeals; 16 honest and 16 deceptive appeals collected from various news and media sites from the United States, the United Kingdom, Canada and New Zealand. This was a larger sample set than is common in high stakes deception studies; some studies have used samples of honest and deceptive behaviour from only one individual (Villar et al., 2011; Vrij & Mann, 2001a), and a body of research has been based on the behaviours of 16 suspects during police interviews (Mann et al., 2002; Mann & Vrij, 2006; Mann et al., 2006).

In all cases classified as deceptive, the appealer was convicted in a criminal court of involvement in the death (or in one case, kidnapping), of their relative. In all cases classified as honest, either another person was convicted of the death of the relative, or the relative was found with no evidence of foul play. In all cases, there was extensive evidence to support the conclusion that the appeals were actually truthful or deceptive. All the appeals were made within a short time frame after the relative went missing or was murdered. For a full description of the sample used in the present study, and the criteria used to determine whether ground truth had been established, see Wright Whelan et al. (2013). As before, the response sheet given to each observer contained one statement for each cue.

Results and Discussion

Again, to assess inter-rater reliability between the five observers, Kendall's W was calculated for each cue. Inter-rater reliability was not significant for six of the cues (p>.05), which were dropped from further analysis (gaze aversion, pretending to cry, no emotional variation, defensive, containing emotion, and credible eye contact). One cue, (no sympathy for the appealer) narrowly missed significance (p = .054), and was retained for further analysis. Average Spearman r for dropped cues ranged from r = .078 (containing emotion) to r = .320 (no emotional variation). Average Spearman r for cues that were retained ranged from r = .350 (no sympathy for the appealer) to r = .584 (high vocal pitch).

The mean of the five observers' scores for each cue in each clip was calculated, so that each clip had a single score for each cue. A MANOVA with Veracity (Truthful / Deceptive) as the between-subjects independent variable, was conducted on the mean cue scores for the cues expected to be related to deception. The results are shown in Table 3. Using Pillai's Trace, the multivariate test yielded a significant result, F (1, 21) = 2.49, p = .038. Follow-up univariate analyses showed significant effects for the following cues, all with

means in the expected direction: fake emotion, fake facial expression, putting on a performance, behaving unnaturally, creepy, dislike for the appealer, and lack of sympathy for the appealer.

Levene's test for homogeneity of variance was significant for the cue fake emotion, and therefore Welch's Test was run on this cue as a check; it produced a result comparable to the univariate test (p = .001).

A similar MANOVA with Veracity as a between-subjects independent variable, was also conducted on the scores for the cues expected to be related to honesty. The results are shown in Table 3. Using Pillai's Trace, the multivariate test showed a result that approached significance, F (1, 19) = 2.10, p = .072; given the result approached significance, and a preponderance of large effect sizes, follow-up univariate analyses were conducted, nonetheless. These showed significant effects for the following cues, all with means in the expected direction: sad eyes, genuinely sad, genuine and heartfelt appeal, urgency, plausible, normal, speaking in a personal and expressive way, voice quivering with genuine emotion, feeling the appealer's pain, and feeling sorry for the appealer.

Levene's test for homogeneity of variance was significant for the cues sad eyes, plausible, and feel sorry for the appealer. Again, Welch's Test produced results comparable to the univariate tests; sad eyes p = .015, feel sorry for the appealer p = .002, plausible p = .005.

Although it does not affect the interpretation of the results and conclusions, it can be noted that in Table 3 the standard deviations for responses to lying appeals are slightly but consistently higher than truthful appeals; however, it appears that this was not a feature of responses to lying and truthfulness per se, as the opposite was the case in the Study 2 (see

Table 2). Further examination of the raw data showed no obvious reasons for these patterns and thus there is no obvious explanation for this anomaly.

General Discussion

In total, 17 cues discriminated between deceptive and honest appealers in the final analyses, and all had large effect sizes in this respect. This endorses the view that using ecologically valid stimulus materials, from a specific context, may elicit specific cues not apparent in typical low stakes, laboratory situations. Indeed, several of the cues identified as discriminating between honesty and deception in appeals are original, notwithstanding that some may be predicted by previous research. Affective responses emerged as a key element in most of the cues which discriminated between deceptive and honest appeals, and the two overarching categories of cues could be described as those relating to 'emotional authenticity', and 'personal reaction to the appealer'.

Cues relating to emotional authenticity (or a lack thereof), are in line with the behavioural control factor of the Four Factor Model (Zuckerman et al., 1981), and the findings suggest not only that emotional authenticity may be useful in discriminating between honesty and deception in this context, but also that untrained observers may be able to detect simulated emotion. The findings concerning fake facial expressions, and sad eyes, are notable in relation to a recent study on appeals by ten Brinke and Porter (2012), in which deceptive appealers were found to produce more expressions of upper face surprise (which the authors suggest was a result of failed attempts to portray sadness), and honest appealers were found to produce more expressions of upper face sadness. Furthermore, the vocal cues are particularly interesting, as for decades efforts have been made to develop voice analysis software to detect deception, using measures such as fundamental frequency, jitter, shimmer, and intensity. However, previous attempts to use measures of vocal characteristics as indicators of deception have generally been without success (see, for example, Giddens et al.,

2013; Horvath, McCloughlan, Weatherman & Slowik, 2013). In contrast, the present results suggest that subjective impressions of vocal characteristics may capture an aspect of deceptive behaviour which it has not been possible to measure more objectively.

The findings relating to personal reactions of observers to appealers are in line with previous research showing that liars are judged to be less pleasant than truth-tellers (DePaulo et al., 2003), and people feel significantly less comfortable when hearing a lie than when hearing a truth (Anderson, DePaulo, Ansfield, Tickle & Green, 1999). A possible explanation for this effect is that, according to the affective response factor of Zuckerman et al.'s Four Factor Model (1981), liars may sound more unpleasant than truth-tellers because the former are more likely to experience, and manifest reactions to, unpleasant negative feelings of guilt and fear. The differences between deceptive and honest appealers in terms of these kinds of cues might also relate to norm violation; it is possible that deceptive behaviour in itself violates certain norms that usually go unnoticed, and, consequently, appears aberrant and unpleasant (hence cue descriptions such as, 'creepy' and 'unnatural'). The more sympathetic personal reactions of the observers to the honest appealers, and their corresponding lack of sympathy for the deceptive appealers, are also of particular interest in relation to recent developments in neuroscience. Research on mirror neurons suggests that we have neural mechanisms which enable a form of direct experiential understanding of what we observe in others, and this permits us to comprehend the emotions of others (Rizzolatti, Fogassi & Gallese, 2006; Singer et al., 2004). Given the present results, therefore, it could be argued that this kind of emotional synchronisation is more likely to occur with genuine emotions than with faked emotions.

The finding that deceptive appealers were rated as being less plausible than honest appealers is consistent with De Paulo et al's (2003) finding that lies are less plausible than truths; however, in the context of an appeal, even 'plausibility', or lack of it, could also be

construed as a reaction to the offender, as it could equally reflect a positive or negative trait attribution about the offender as a person. This brings us to a more general point about the cues identified in the present study; although some of the cues are accommodated by existing theoretical approaches, the majority of them are not. It might, therefore, be useful to entertain some other, newer perspectives on cues to deception. As Salfatti (2003) says, 'Homicide grows out of a transaction between individuals. This transaction is a product of the individuals and their relationship' (p.499). Hence, one possible alternative explanatory approach is the 'social interactionist' (SI) perspective suggested by Wright Whelan et al. (2013). According to the SI approach, some cues to deception may reflect not the act of deception itself, but the nature of the social relationships between the offender and the victim in contexts where this is relevant. So, for example, one might expect honest appealers to appear more genuinely sad (as an overall impression, facially, and vocally), than somebody who has recently killed his or her relative, was presumably motivated to do so, and may need to simulate these emotions. Moreover, one might expect honest appealers to display more urgency in their attempt to find out what happened to their relative, than deceptive appealers who clearly do not have the same motivation regarding their relative. Importantly, this SI approach accommodates some of the more context-specific findings reported by ten Brinke and Porter (2012), in their analysis of appeals. The researchers found that deceptive appealers were more likely to produce partial facial expressions of disgust and happiness, and less likely to produce facial expressions of sadness/distress. According to the SI approach, one would expect appealers who had actually killed their relative to be more likely than honest appealers to leak expressions of disgust when talking about their relative, or to leak expressions of happiness when talking about the absence of their relative, and honest appealers to exhibit more expressions of sadness/distress.

The SI approach is also useful in explaining newly identified, context specific cues in a different high stakes situation. As mentioned previously, several such cues were identified in Harpster et al.'s (2009) study of 911 homicide calls, in which, for example, honest callers were more likely to make a plea for help for the victim; according to the SI approach, one might expect somebody who has killed a victim to be less likely to plea for help for that victim. Deceptive callers, who had just killed the victim, were more likely to insult the victim, and to describe the death of the victim in terms of a problem, and to accept that the victim was dead; again, these behaviours could be regarded as reflecting the nature of the social interaction between the offender and the victim (in the case of acceptance of death, this may be a cognitively driven behaviour reflecting a change in the status of the victim in the relationship).

A second, and related, way of interpreting some of these cues in the present cases is in terms of the 'individual behavioural profile' (IBP) account of deception (Wright Whelan et al., 2013). In the context of appeals, and in many high stake situations, honest and deceptive individuals differ not only in terms of whether they are lying, but also whether they have committed a serious criminal act. According to the IBP approach, some of the cues identified in the present studies may be related to characteristics of individuals who engage in aberrant behaviour (kill their relative and then lie about it), rather than actual acts of deception. For example, as with the act of lying, less genuine emotion and empathy towards the victim are consistent with the stereotype of a psychopath (Davies & Feldman, 1981). Also 'behaving unnaturally', looking 'creepy', and not appearing likeable or worthy of sympathy, might reflect characteristics of people who murder relatives, rather than behaviours related to the act of deception. Cues relating to general characteristics of this kind will clearly not emerge as discriminatory in low stakes laboratory experiments, and because they may not necessarily relate to the act of deception per se, they may not be apparent in within subjects comparisons

of truthful and deceptive conditions, even in high stakes situations. Consequently, while, at a practical level, real-life high stakes situations might potentially provide a useful source of cues for making probabilistic estimates about who is lying and who is not, they may pose some methodological problems for researchers attempting to determine cues to deception, as distinct from cues that identify more global characteristics of offenders.

Clearly, as the SI and IBP approaches are newly developed, their utility in predicting and explaining behaviours related to deception or to honesty has not been firmly established. It is hoped, however, that they may be expanded to provide a theoretical rationale for cues that emerge only in high stakes situations, and are not readily accommodated by existing theoretical approaches. It is likely that the SI and IBP approaches will have relevance beyond the context of appeals, and may be particularly relevant to contexts which involve violent crimes in which there are offenders and victims. However, what these specific contexts may be requires further investigation.

The present studies were obviously limited in scope, both in terms of the sample size and the range of robust statistically significant results, and consequently can only be considered exploratory and tentative at this stage. Nevertheless, several previously uninvestigated cues were identified, and the overlap in findings between Studies 2 and 3 for cues found to be significant in Study 3 (all with substantial effect sizes) suggests that although the cues were not generated from existing theoretical approaches, they may have utility in the context of appeals. As the research was based on field data, clearly key experimental elements are missing. However, it was considered that this limitation was outweighed by the need for ecologically valid data; hundreds of laboratory-based experiments on deception have already been conducted, but replicating a high stakes environment is likely to be impossible in the laboratory. The identification of previously uninvestigated cues underlines the importance of using ecologically valid stimulus materials

in deception research. Notwithstanding this, the western bias of the stimulus materials used in the present studies is acknowledged, and it is certainly possible that some behaviours may not generalise across cultures. A further limitation of the stimulus materials used was that the scope of the image varied in the appeals, so that in some cases just the head was shown, and in others more of the body was shown; it is possible that this may have an effect on observer responses.

In terms of future research, initially it would be beneficial to replicate the findings of the present studies on a new sample of appeals, and especially to include appeals from non-western cultures. Furthermore, the extent to which some of the newly identified cues may have utility in different high stakes contexts requires investigation; relatedly, the factors which may underlie the production of the cues (such as emotional authenticity, norm violation, and the personal reaction elicited in observers) may have relevance across particular high stakes situations and again, this remains to be investigated in future research. It is also hoped that the utility of using a data-driven, context-specific, bottom-up approach has been demonstrated, and that this approach may be adopted in the examination of other high stakes situations.

The basic premise of the research presented here was that the use of real life examples of high stakes honest and deceptive behaviour as stimulus materials, would result in findings different to those typically produced in low stakes deception research. This premise was supported, as a large number of behaviours were found to differ between honest and deceptive appeals, all with effect sizes much larger than is common in deception research, and it is hoped that this will reinforce the importance of ecological validity in deception research. In addition, the identification of previously uninvestigated cues suggests that it may be important to consider context, a factor which is often overlooked in deception research. Furthermore, it is hoped that the development of new theoretical approaches as suggested

above may be beneficial when considering real world deceptive behaviour. It is unlikely that a fool proof system of detecting deception will be established in the foreseeable future (if ever); however, it is hoped that the present findings suggest fruitful avenues for future research.

Author Notes

Clea Wright Whelan is a lecturer at the University of Chester. Her research interests encompass behavioural correlates of deception, deception detection, and observer decision making processes when making credibility judgements.

Graham F. Wagstaff is Professor of Cognitive Social Psychology at Liverpool University. He has specialist research interests in the areas of hypnosis, forensic interviewing, eyewitness testimony and the psychology of justice.

Jacqueline Wheatcroft is in the Department of Psychological Sciences, Witness Research Group at the University of Liverpool. Her current research interests are forensic interviewing of witnesses, questioning protocols in legal contexts, memory enhancement tools (i.e. Liverpool Interview Protocol) and relationships between confidence and accuracy.

References

Adams, S. H., & Jarvis, J. P. (2006). Indicators of veracity and deception: an analysis of written statements made to police. *The International Journal of Speech, Language and the Law, 13* (1), p.1-22

Anderson, D. E., DePaulo, B. M., Ansfield, M. E., Tickle, J. J. & Green, E. (1999). Beliefs about cues to deception: Mindless stereotypes or untapped wisdom? *Journal of Nonverbal Behaviour, 23,* 67-89

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3,* 77-101. DOI: 10.1191/1478088706qp063oa

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum

Davis, M., Markus, K. A., Walters, S. B., Vorus, N. & Connors, B. (2005). Behavioral cues to deception vs. topic incriminating potential in criminal confessions. *Law and Human Behavior, 29* (6), p.683-704. DOI: 10.1007/s10979-005-7370-z

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K. and Cooper, H. (2003). Cues to Deception. *Psychological Bulletin, 129* (1), 74-118. DOI: 10.1037/0033-2909.129.1.74

DePaulo, B. M. & Morris, W. L. (2004). Discerning lies from truths: behavioural cues to deception and the indirect pathway of intuition. In Granhag, P. A. & Stromwall, L. A. (Eds), *The Detection of Deception in Forensic Contexts* (pp.15-40). Cambridge: Cambridge University Press

Frank, M. G. & Svieteva, E. (2012). Lies worth catching involve both emotion and cognition: Commentary on Vrij and Granhag (2012). *Journal of Applied Research in Memory and Cognition, 1* (2), 131-133. DOI: 10.1016/j.jarmac.2012.04.006

Giddens, C. L., Barron, K. W., Byrd-Craven, J., Clark, K. F. & Scott Winter, A. (2013). Vocal indices of stress: A review. *Journal of Voice.* Advance online publication. doi:10.1016/j.jvoice.2012.12.010

Glass, G. V. & Stanley, J. C. (1970). *Statistical Methods in Education and Psychology.* Englewood Cliffs, N.J.: Prentice-Hall

Granhag, P. A. & Stromwall, L. A. (Eds), *The Detection of Deception in Forensic Contexts* (pp.15-40). Cambridge: Cambridge University Press

Harpster, T., Adams, S. H., and Jarvis, J. P. (2009). Analyzing 911 homicide calls for indicators of guilt or innocence. *Homicide Studies, 13* (1), 69-93

Hayes, N. (2000). *Doing Psychological Research.* Buckingham, England: Open University Press

Horvath, F., McCloughlan, J., Weatherman, D. & Slowik, S. (2013). The accuracy of auditors' and layered voice analysis operators' judgements of truth and deception during police questioning. *Journal of Forensic Sciences.* Advance online publication. doi:10.1111/1556-4029.12066

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury Thomson Learning, Inc.

Koper, R. J. & Sahlman, J. M. (1991). The behavioral correlates of real-world deceptive communication. Paper presented at the Annual Meeting of the International Communication Association. Retrieved online 13 March 2009, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/24/18/ae.pdf

Levine, T. R., Anders, L. N., Banas, J., Baum, K. L., Endo, K., Hu, A. D. S. & Wong, N. C. H.

(2000). Norms, expectations and deception: A norm violation model of veracity judgements. *Communication Monographs, 67* (2), 123-137

Mann, S. & Vrij, A. (2006). Police officers' judgement of veracity, tenseness, cognitive load and attempted behavioral control in real-life police interviews. *Psychology, crime & Law, 12* (3), 307-319. DOI: 10.1080/10683160600558444

Mann, S., Vrij, A. & Bull, R. (2002). Suspects, lies and videotape. An analysis of authentic, high-stakes liars. *Law and Human Behaviour, 26,* 365-376

Mann, S., Vrij, A. & Bull, R. (2006). Looking through the eyes of an accurate lie detector. *The Journal of Credibility Assessment and Witness Psychology, 7* (1), 1-16

O'Sullivan, M. & Ekman, P. (2004). The wizards of deception detection. In Granhag, P. A. & Stromwall, L. A. (Eds), *The Detection of Deception in Forensic Contexts* (pp.269-286). Cambridge: Cambridge University Press

Porter, S. & ten Brinke, L. (2010). The truth about lies: What works in detecting high-stakes deception? *Legal and Criminological Psychology, 15,* 57-75. DOI: 10.1348/135532509X433151

Rizzolatti, G., Fogassi, L. & Gallese, V. (2006). Mirrors in the mind. *Scientific American, 295* (5), 54-61

Salfatti, C. G. (2003). Offender interaction with victims in homicide: A multidimensional analysis of frequencies in crime scene behaviors. *Journal of Interpersonal Violence*. 18, 499-512.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J. & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303* (5661), 1157-1162

Tabachnick, B. G. & Fidell, L. S. (2001). *Using Multivariate Analysis.* Boston: Allyn & Bacon

ten Brinke, L., & Porter, S. (2012). Cry me a river: Identifying the behavioral consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior* (in press)

Villar, G., Arciuli, J., & Mallard, D. (2011). Use of "um" in the deceptive speech of a convicted murderer. *Applied Psycholinguistics, 33* (1), 83-95. doi: 10.1017/S0142716411000117

Vrij, A. & Mann, S. (2001a). Telling and detecting lies in a high-stake situation: the case of a convicted murderer. *Applied Cognitive Psychology, 15* (2), 187-203

Vrij, A. & Mann, S. (2001b). Who killed my relative? Police officers' ability to detect real-life, high-stakes lies. *Psychology, Crime and Law, 7,* 119-132

Wright Whelan, C. (2009). *The art of deception and the science of detection: Cues to deceit in real life, high-stakes situations* (Unpublished Master's dissertation). University of Liverpool, Liverpool, United Kingdom

Wright Whelan, C., Wagstaff, G. F. & Wheatcroft, J. M. (2032). High Stakes Lies: Verbal and Nonverbal Cues to Deception in Public Appeals for Help with Missing or Murdered Relatives. *Psychiatry, Psychology and Law*, published online 23 September 2013. DOI: 10.1080/13218719.2013.839931

Zuckerman, M., DePaulo, B.M & Rosenthal, R. (1981). Verbal and nonverbal communication in deception. In L. Berkowitz (Ed), *Advances in Experimental Social Psychology* (Vol. 14, pp.1-59). New York: Academic Press

Table 1: *Study 1 qualitative analysis; key themes with composite categories. Number of units relating to each category is indicated in brackets).*

Deceptive appealers

Fake emotion
Fake emotion (11)
Fake facial expression (4)
Putting on a performance (6)
Pretending to cry (6)
High vocal pitch (15)
No emotional variation (4)

Lack of emotion
Lack of emotion (6)
Painless recall (7)
Calm/calculated (9)
Lack of emotion in voice (4)
Emotionally cold (7)
Lack of facial expression (4)
Brutal language/detail (3)

Distancing
Impersonal (7)
Focus on others (9)
First person pronouns (2)
Equivocation (3)
Irrelevant/generic statements (2)
Not focused on relative (8)
Lists (7)
Gaze aversion (8)

Implausibility
Nonsensical (9)
Making assumptions (2)
Lack of hope of finding missing relative alive (7)
Indication of dispute (2)
Defensive (3)
Fear/nervousness (9)

Negative personal reaction
Dislike for the appealer(6)
Creepy (2)
Behaving unnaturally (3)
No sympathy for the appealer (6)

Honest appealers

Genuine sadness
Genuine sadness (8)
Facial expressions (4)
Sad eyes (5)
Desperation/urgency (4)

Genuine/heartfelt
Genuine/heartfelt (7)
No prepared speech (5)
Positive emotion expressed towards relative (6)
Smile appropriate to speech content (2)
Genuine behaviours (3)
Credible eye contact (6)

Containing emotion
Containing emotion (15)
Calmness appropriate to speech content (9)
Trying to be sensible (8)
Hopeful of finding missing relative (2)
Avoiding brutal language/detail (2)

Sympathetic personal reaction
Feel the appealers' pain (3)
Feel sorry for the appealer (7)
Normal (3)
Behaving as expected (4)

Table 2: *Kendalls's W, Means, SDs, F ratios, and effect sizes for cues which discriminated between deceptive and honest appealers, or had large effect sizes in the expected direction, in Study 2*

	Kendall's W (all values $p < .05$)	Deceptive Mean (SD)	Truthful Mean (SD)	*F* ratio *df* (1,7)	Effect size (*d*)
Cues to deception					
Fake emotion*	.61	3.66 (0.32)	2.48 (0.84)	$F = 6.90$	1.86
Fake facial expression*	.44	3.46 (0.24)	2.56 (0.64)	$F = 7.04$	1.86
Putting on a performance*	.61	3.98 (0.46)	2.72 (0.91)	$F = 6.16$	1.73
Creepy*	.41	3.54 (0.37)	2.24 (0.74)	$F = 10.08$	2.22
No sympathy for appealer*	.58	3.71 (0.36)	2.48 (0.92)	$F = 6.22$	1.76
High vocal pitch*	.53	3.25 (0.91)	2.08 (0.30)	$F = 7.40$	1.73
Gaze aversion*	.65	4.13 (0.49)	2.60 (0.93)	$F = 8.72$	2.04
Lists*	.48	3.71 (0.51)	2.56 (0.78)	$F = 6.41$	1.75
Behaving unnaturally	.61	3.63 (0.62)	2.88 (1.09)	$F = 1.46$	0.83
Dislike for the appealer	.48	3.49 (0.36)	2.64 (0.97)	$F = 2.67$	1.16
Pretending to cry	.73	3.16 (1.16)	2.04 (0.61)	$F = 3.56$	1.21
No emotional variation	.43	3.63 (0.42)	3.00 (0.68)	$F = 2.57$	1.10
Emotionally cold	.43	3.54 (0.37)	2.60 (0.97)	$F = 3.28$	1.28
Equivocal language	.42	2.79 (0.17)	2.20 (0.57)	$F = 3.94$	1.40
Focus on others	.57	3.45 (0.53)	2.72 (0.94)	$F = 1.89$	0.96
Defensive	.46	2.89 (0.27)	2.20 (0.63)	$F = 4.06$	1.42
Does not make sense	.61	3.24 (0.46)	2.40 (0.93)	$F = 2.68$	1.14
Cues to honesty					
Sad eyes*	.58	2.19 (0.41)	3.22 (0.85)	$F = 7.15$	1.54
Containing emotion*	.43	2.40 (0.59)	3.60 (0.42)	$F = 12.73$	2.34
Genuine facial expression*	.51	2.49 (0.24)	3.44 (0.75)	$F = 5.78$	1.71
Credible eye contact	.61	2.19 (0.41)	3.32 (0.89)	$F = 5.44$	1.63
Genuinely sad	.61	2.40 (0.49)	3.40 (0.76)	$F = 5.12$	1.56
Genuine/heartfelt appeal	.58	2.29 (0.56)	3.24 (1.06)	$F = 2.58$	1.12
Feel sorry for the appealer	.53	2.49 (0.24)	3.32 (0.88)	$F = 3.31$	1.29
Normal	.50	2.23 (0.39)	3.04 (0.92)	$F = 2.69$	1.15
Feel the appealer's pain	.51	2.49 (0.33)	3.20 (0.91)	$F = 2.19$	1.04
Genuine behaviours	.57	2.25 (0.57)	3.24 (0.80)	$F = 4.26$	1.43

* $p < .05$

Table 3: *Kendall's W, Means, SDs, F ratios, and effect sizes for cues which discriminated significantly between deceptive and honest appealers in Study 3*

	Kendall's W (all values $p < .05$)	Lying Mean (SD)	Truthful Mean (SD)	*F* ratio *df* (1,31) All values $p < .05$	Effect size (*d*)
Cues to deception					
Fake emotion	.55	3.33 (0.79)	2.49 (0.48)	$F = 13.50$	1.29
Fake facial expression	.58	3.35 (0.73)	2.49 (0.60)	$F = 13.25$	1.29
Putting on a performance	.50	3.38 (0.71)	2.61 (0.52)	$F = 11.92$	1.24
Behaving unnaturally	.50	3.40 (0.70)	2.76 (0.63)	$F = 7.38$	0.96
Creepy	.53	3.30 (0.71)	2.60 (0.47)	$F = 10.61$	1.16
Dislike for the appealer	.54	3.42 (0.65)	2.74 (0.46)	$F = 11.87$	1.21
No sympathy for appealer	.48	3.24 (0.65)	2.56 (0.41)	$F = 11.96$	1.25
Cues to honesty					
Sad eyes	.60	2.55 (1.05)	3.33 (0.53)	$F = 6.89$	0.94
Genuinely sad	.61	2.65 (0.85)	3.51 (0.57)	$F = 11.35$	1.19
Genuine/heartfelt appeal	.66	2.55 (0.87)	3.40 (0.63)	$F = 9.94$	1.12
Urgency	.57	3.00 (0.83)	3.65 (0.77)	$F = 5.27$	0.81
Plausible	.51	2.71 (0.68)	3.30 (0.36)	$F = 9.42$	1.08
Normal	.49	2.74 (0.64)	3.35 (0.44)	$F = 10.03$	1.11
Personal & expressive voice	.49	2.76 (0.65)	3.40 (0.63)	$F = 7.86$	1.00
Voice quivering with emotion	.56	2.50 (0.81)	3.41 (0.74)	$F = 11.00$	1.17
Feel the appealer's pain	.56	2.70 (0.75)	3.35 (0.58)	$F = 7.37$	0.97
Feel sorry for the appealer	.57	2.68 (0.78)	3.48 (0.51)	$F = 11.80$	1.21